

SOCIAL INDICATORS 1973: ¹
Statistical Considerations

by

Stephen E. Fienberg² and Leo A. Goodman³

Technical Report 229

July 1974

¹This paper is a revised version of remarks prepared for the Review Symposium on Social Indicators: 1973, sponsored by the SSRC Center for Coordination of Research on Social Indicators, Washington, D. C., February 21-23, 1974. It is based on conversations and written memoranda from several of our colleagues, as well as on our own conversations and observations. Specifically, we wish to thank Albert Biderman, O. Dudley Duncan, William Kruskal, Robert Parke, Jr., Joseph Waksberg, and Hans Zeisel for their help. In addition, we have drawn upon such sources as Federal Statistics: The Report of the President's Commission, Volumes I and II (1971), U. S. Government Printing Office, Washington, and Mathematical Sciences and Social Sciences (prepared in connection with the Survey of the Behavior and Social Sciences), edited by William Kruskal (1970), Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

²Department of Applied Statistics, School of Statistics, University of Minnesota, St. Paul, Minnesota.

³Departments of Statistics and Sociology, University of Chicago, Chicago, Illinois. This work was supported in part by Research Contract No. NSF 31967X from the Division of the Social Sciences of the National Science Foundation to the Department of Statistics, University of Chicago.

Introduction

A reasonable scenario for the development of a system of social indicators would consist of several stages. A logical starting place is the identification of broad areas of social concern (e.g., public safety might be one such area. Once these areas have been identified, as the second step in the scenario, we can focus our attention on specific phenomena that pertain to each area of concern, (e.g., the incidence of criminal behavior and the risk of victimization); these phenomena, if properly measured by ideal indicators (e.g., by crime and victimization rates), would reflect or elucidate the state of society and the changes (if any) that are taking place in it. Then as the third step, we can compare the ideal indicators with the ones available (e.g., the rate of offenses reported to the police as given in the FBI Uniform Crime Reports), with an eye toward accepting substitutes for some ideal measures, and possibly insisting on the development of new measures (e.g., victimization rates from national and local surveys). Finally, we must decide how often to make measurements and how best to present or report the measurements we make.

Although the statistician in this scenario might focus on methods of data collection, on assessing and maintaining the quality of measurements, and on the manner in which they are reported, we choose to view the statistician's role more broadly, especially since it is difficult for us to separate the quality of measurements from what it is we actually would like to measure. Thus our comments on Social Indicators, 1973 will be concerned not only with data collection and sampling techniques, and with accuracy and reporting, but also with broader issues related to the social indicators enterprise.

One final preliminary comment is in order. Some of the remarks we shall make in the paper are critical in the sense that we suggest a variety of ways in which Social Indicators can be improved upon statistically. It is our belief that many of our criticisms are equally applicable (if not more so) to social indicator reports produced by other countries, such as France, Germany, and the United Kingdom. Moreover, in no way should our remarks be interpreted as being critical of the many achievements of the federal statistical system which are reflected in Social Indicators, 1973.

One final preliminary comment is in order. Some of the remarks we shall make in the paper are critical in the sense that we suggest a variety of ways in which Social Indicators can be improved upon statistically. It is our belief that many of our criticisms are equally applicable (if not more so) to social indicator reports produced by other countries, such as France, Germany, and the United Kingdom. Moreover, in no way should our remarks be interpreted as being critical of the many achievements of the federal statistical system which are reflected in Social Indicators, 1973.

Some Reporting Problems

When preparing a social report such as Social Indicators, 1973, one faces at least the following kinds of statistical problems:

- (a) How to abstract a salient summary of major ongoing studies, especially those with complex statistical analyses?
- (b) How to check whether the abstract distorts or disguises important features of a more complete presentation?
- (c) How to check whether the abstract reproduces uncritically inadequacies in the original sources - - inadequacies in the reporting or in the related original analyses?

There is no prescription we can offer that tells how to solve these problems, but the following example may serve to illustrate them.

Charts 3/8 through 3/11 in the Education chapter [;82-85] purport to show "which groups in the population are behind or ahead and by how much," based on the results presented by the National Assessment of Educational Progress. The National Assessment data come from a national sample of 17 year-olds, selected under the direction of professional statisticians using intricate sampling techniques. The analyses of these data were performed under the supervision of distinguished statisticians, and the analyses and graphical presentations in the National Assessment reports issued to date reflect this professional strength. Yet the information from the National Assessment project that is abstracted in Social Indicators, 1973 can easily be misunderstood.

The baseline for Charts 3/8 through 3/11 in each case is the median performance of all 17 year olds, and each bar-graph illustrates median differences between or among groups, over all questions in the achievement test. What is misleading about these graphs? First, the scales from graph to graph are somewhat different. Second, the apparent apportionment of differences to groups is a function mainly of group size. To see this, we focus on White-Black differences in Science Achievement (see Figure 1), and for simplicity we work with means rather than medians.⁴ We let

\bar{X}_W = average for Whites; N_W = number of Whites

\bar{X}_B = average for Blacks; N_B = number of Blacks

\bar{X} = average for total population; $N = N_W + N_B$ = number in total population.

Then

$$N\bar{X} = N_B\bar{X}_B + N_W\bar{X}_W,$$

and we find that

$$\bar{X}_W - \bar{X} = \frac{N_B}{N} (\bar{X}_W - \bar{X}_B)$$

$$\bar{X}_B - \bar{X} = \frac{N_W}{N} (\bar{X}_B - \bar{X}_W).$$

Thus each of the bars in the two group comparisons simply represents the appropriate group difference (i.e., $\pm (\bar{X}_W - \bar{X}_B)$), magnified by the relative size of the other group. In this case, since the Blacks represent a smaller proportion of the population relative to the Whites, the graphs appear to show them as being more "disadvantaged." If Blacks had represented, say, 87 percent of the population and Whites 13 percent rather than the reverse, the overall average \bar{X} would shift towards the Black average, and the picture

⁴The comments we make are applicable in a slightly revised form to medians.

for the deviations would look dramatically different (see Figure 2). In this case Blacks would appear only modestly disadvantaged while Whites would become more "advantaged." Thus the relative magnitude of each bar reflects only the relative size of the other group.

We note that most of the other graphs in the National Assessment report include the equivalents of \bar{X} , \bar{X}_W , and \bar{X}_B , and not just the deviations, $\bar{X}_W - \bar{X}$ and $\bar{X}_B - \bar{X}$. Thus these graphs are not subject to the above criticism.

These graphical displays may easily be elaborated, to take advantage of the beautiful colors and graphics of Social Indicators, 1973. What we have here are displays of "typical" population values for Blacks and for Whites. But we can also present information bearing on the population distribution of achievement by using bands to illustrate quartiles, or other suitable quantiles. In Figure 3 we give an example of such bands for Science Achievement, using hypothetical quartiles. These bands represent one point in time. How interesting they would be if used with the time series data that the National Assessment project will be collecting over the next few years. The purpose of plotting these bands is to introduce the variability associated with the population of achievement scores.

In addition to the variability considered above, there is another type of variability that we must also consider; viz., the variability or error which results from estimating (from a sample) the typical values for a population. Later in this paper, we turn to this notion of variability and the ways we feel it should be considered.

Our earlier discussion of Charts 3/8 through 3/11 served to illustrate problems (a) and (b) as described at the beginning of this section. Our

next example will serve as another illustration of problem (b). We consider now the display of grade enrollment by race, sex, and age in Chart 3/7 [; 80]. The graph presents the percent of students below the modal grade. A more symmetric picture would have presented the percent in the modal grade and the percent above the modal grade as well as the percent below it. Although the percent below the modal grade is of special interest, information about the complete distribution may enable us to gain greater understanding of the meaning of the magnitudes displayed. The information presented in Chart 3/7 in its present form is potentially misleading.

Accuracy and Error Structure

How accurate are the data and the series reported in Social Indicators, 1973? The Introduction does touch on accuracy as follows: "Most of the series included have been taken from Federal sources; their quality can be verified by those agencies. For data compiled by non-Federal sources, we have, wherever possible, relied on the judgment of those working directly with the data regarding their suitability for this publication" [; xiv]. Yet in most cases Federal agencies have not prepared adequate studies of error for their series. (The Bureau of the Census is one of the important exceptions in this regard.) Thus agency verification can usually be only unsupported assertion. It is unfortunate that the Statistical Policy Division of OMB has produced a major statistical publication without a serious discussion of error, especially in light of the recommendations and the general thrust of the President's Commission of Federal Statistics.⁵

The absence of a discussion of data accuracy seems unfortunate as a matter of principle and statistical standards, and it may also lead to misunderstandings and mistakes. For example, the relatively innocent reader may note a difference between two tabulated values dominated by random variation and conclude that some real pattern exists when in fact this is not the case.

We are all familiar with the notion of random sampling error; in order to keep such errors at low levels, it is necessary for us to have samples that are appropriately large and well-designed. While many of the Charts in Social Indicators, 1973 are based on sample data, few of the related Technical Notes give any details on sample design, and in only one case are we given

⁵ See Federal Statistics: The Report of the President's Commission, Volume I (1971), U. S. Government Printing Office, Washington, D. C., pp. 29-33.

an estimate of sampling error.⁶ In addition to the random component of sampling error, whenever nonprobability sampling techniques are used at some level of a survey, there enters the possibility of systematic sampling errors.

Whether or not sampling error is present in a study, there may be appreciable errors of measurement, such as nonresponse and various types of response errors (misunderstandings, failures of memory, deliberate falsehoods). These measurement errors may be random or systematic, and the taking of a census rather than a sample does not remove such systematic errors. For example, it is well-known and well-documented that the U. S. Census of Population has a systematic undercount (bias), the magnitude of which differs by race. For the 1970 census, it has been estimated that approximately 1.9 percent of the whites and 7.7 percent of the non-whites were not counted, and that the undercount of some non-white male five-year age groups was as high as 18.5 percent.⁷ (These are estimated rates, however, and they too are subject to various sources of error.) This systematic measurement error is especially relevant for the Population chapter of Social Indicators, 1973, where unadjusted census figures are used for several charts, but it is also relevant for data presented in other chapters as well. For example, if adjustments to reported crime rates took the above mentioned undercount into account, the rates thus obtained might present a somewhat different picture from that associated with the unadjusted rates presented in Social Indicators, 1973.

Some examples of systematic measurement error in population data

⁶See the note on p. 63 for Charts 2/20 and 2/21.

⁷Siegel, J. S. (1974). "Estimates of Coverage of the Population by Sex, Race, and Age in the 1970 Census." Demography 11, 1-23.

have been noted in Social Indicators, 1973. For example, in the Housing chapter, the Technical Notes [; 202-203] point out that the 1960 Census of Housing underestimated the number of substandard units in the U. S. by 536,000 or about 6 percent of all substandard occupied housing units. When these data are disaggregated by social and demographic classifications, however, no adjustments for this bias are made, even though the disaggregated data are susceptible to this bias. (An appropriate adjustment for each disaggregated class would, of course, be preferable, but if the information that would be needed to do this is not available, then a uniform adjustment for the disaggregated classes would be better than no adjustment.) Nevertheless, the discussion of the effects of measurement error with regard to this particular index in Social Indicators, 1973 is good, and we would like to see other thoughtful discussions of this sort.

The random component of measurement error often results from the imperfection of the measuring device, and the simplest way to get a handle on the magnitude of this measurement error is to repeat measurements independently and/or to compare measurements obtained using the usual measuring device with measurements of the same quantity obtained using a "more accurate" measuring device. Neither of these approaches is necessarily easy to carry out, and nowhere in Social Indicators, 1973 is there a discussion of this random component of measurement error or of attempts by various agencies to measure it.

Many of the indices used in Social Indicators, 1973, both those used directly and those used indirectly in the construction of other indices reported in the volume, are subject to several sources and types of error.

For example, the Consumer Price Index (CPI), which is used to adjust figures for across-time comparisons in the Income chapter, is based on a complex network of samples, not all of which are probabilistic. By instituting a replication design the Bureau of Labor Statistics has attempted to measure both types of random error (i.e., sampling and measurement error), but we have not been able to find any details on systematic errors, nor on how the errors may vary over time.⁸ Furthermore, the Commissioner of the Bureau,

Julius Shiskin, recently noted that:

The weighting of the CPI to take account of the proportion of disposable income spent in various items in the index, is based on a survey of consumer price patterns in 1960-61. A new survey is getting under way,⁹ but the results will not be available until 1977.

Whether much or little is known about the error structure of a survey or a particular index, it is not sufficient to refer those wishing to examine in greater detail the material in Social Indicators, 1973 to the "quoted sources," many of which are unpublished reports or studies. Even the President's Commission on Federal Statistics was unable to obtain detailed information on error structure from a large number of federal agencies. The Commission's Report notes that "although there was considerable variation, both for different statistics

⁸ For various statistical discussions of the CPI see:

- (a) Kruskal, W. H. and Telser, L. G. (1960). "Food Prices and the Bureau of Labor Statistics," J. Bus. Univ. Chicago 33, 258-285.
- (b) BLS Handbook of Methods for Surveys and Studies (1971). Bulletin 1711. U. S. Government Printing Office, Washington, D. C.
- (c) Griliches, Zvi (ed.). (1971). Price Indexes and Quality Change. Harvard University Press. See especially pp. 185-197 and pp. 233-234.

⁹ Excerpts of a statement by J. Shiskin presented at the Washington Journalism Center, as reported in the Minneapolis Tribune, February 8, 1974.

in the same agency and across agencies, the responses to the [Commission's] survey showed disappointingly little knowledge of error structure.

Sampling errors were estimated for most statistics based on probability samples, but there were, with only a few exceptions, very few analyses of response and other nonsampling errors, even in cases in which, because of long recall or the use of incomplete records, these were likely to be substantial."¹⁰

It is our hope that in conjunction with future editions of Social Indicators, the Statistical Policy Division of OMB will compile in a form suitable for publication detailed information regarding what is known and what is unknown about the error structure for each of the series in the main report. This information should include descriptions (where relevant) of:

- (a) sampling frame, sampling plan, and (effective) sample size,
- (b) estimates of sampling error,
- (c) any special or nonstandard aspects of questionnaire design or interview procedures,
- (d) non-response rates, and treatment of missing observations (if the problem is substantial),
- (e) degree of consistency and compatibility with related series of measures.

These descriptions need not be voluminous; they need only be summary in nature, with references to more detailed technical presentations.

¹⁰ Federal Statistics: The Report of the President's Commission. Vol. II, (1971), U. S. Government Printing Office, Washington, D. C. "How much do agencies know about error structure," by H. Grubert, pp. 297-334.

This information on error structure, required by those who wish to draw inferences from the data in Social Indicators, should probably be published as a companion booklet, and only brief statements should be included in the Technical Notes in the main report. If it is unrealistic to expect such a companion booklet, then a brief checklist format could be included in the Technical Notes which would alert the reader to what is known or unknown about the error structure for each data set or chart. The Bureau of the Census, for example, provided the President's Commission on Federal Statistics with a detailed eight-page summary of information on the error structure of the Current Population Survey (CPS), and data from the CPS have been used in Social Indicators, 1973 in the chapters on Employment, Income, Housing and Population. Something less detailed would suffice for the next edition of Social Indicators.

Objective vs. Subjective

Much discussion in the field of social indicators has focused on the use of objective versus subjective measures, and the Introduction of Social Indicators, 1973 touches on this point [; xiii]. We believe that it is important to point out that there are two different senses in which a measure can be objective or subjective.

William Kruskal notes that phenomena may be subjective in the sense of being inside people's heads (attitudes, aspirations, happiness) or objective in the sense of being directly observable (dead-alive). Similarly, modes of measurement may be subjective (opinions about the magnitude of the crime problem in 1974) or objective (actual counts of reported crimes in various categories). Of course, there are philosophical difficulties as to what is objective and what is subjective, and there are intermediate, blurry cases, but roughly speaking, we can think in terms of the 2 x 2 cross-classification:

		Mode of Measurement	
		Subjective	Objective
Phenomenon of Interest	Subjective	a	b
	Objective	c	d

In reading through Social Indicators, 1973, we have found examples of indicators that correspond to each of the four cells. Of course, the bulk of the measurements appear to be objective-objective, but a closer examination reveals that the mode of measurement in many of the so-called objective-objective cases might be classified as subjective. For example, are self-reports, such as those used in the compilation of disability data in the Health chapter objective or subjective? This depends in fact on instructions given to respondents, and we are not provided with these

details in the Technical Notes. Indeed, the classification of various sets of data according to the 2 x 2 table described above is itself subjective in character.

Now let us turn to the other three cells in this 2 x 2 table (i.e., the a, b, and c cells). The indices in the Housing chapter of substandard housing and crowded conditions are examples of subjective phenomena that are measured objectively (see Charts 6/1 through 6/7). The notion of "substandard" is a subjective one, but the Bureau of the Census has used objective criteria to measure these phenomena. Such measures are often of interest to social scientists because they are related to social concerns and are available on a consistent basis over time. In the present case, the index of substandard housing has been criticized, in part, as a consequence of its consistency over time. The definition has been consistent, but many complain that it is no longer meaningful.

Albert Biderman suggests another example of subjective phenomenon-objective measurement: the Uniform Crime Reports offense series, which he claims serve as important and quite accurate state-of-society indicators as they reflect people's perceptions of the magnitude of the crime problem. He would thus dispute the common view of these series as objective-objective.

Chart 1/26 provides us with an example of a subjective measure of a phenomenon that appears, at first glance, to be objective. This chart is based upon the respondent's assessment of the confidence he has with respect to his access to "good" medical care. If "good" had been

defined for the respondents, then the phenomenon might be objective; if "good" were not defined then we would have the respondent's perception of what is "good," in which case the phenomenon would be subjective. The Technical Notes do not provide us with enough information to determine how "good" was actually defined in the study.

Far more prevalent than subjective-objective and objective-subjective indices are subjective-subjective indices in Social Indicators, 1973. For example, in the Employment chapter, there is a subjective evaluation of the highly subjective notion of "job satisfaction" (see Charts 4/17 and 4/16 [; 123-124]). This is a reasonable approach to one dimension of a quality of life index. Further examples of subjective-subjective measures can be found in the chapters on Public Safety [; 58-59] and Housing [; 200-201].

Statistical Analysis and Interpretation

One of the more striking features of Social Indicators 1973 is the apparent lack of analysis and interpretation of the statistics included. Obviously, judgment has been exercised in the choice of which indicators to include, which to eliminate, what types of disaggregation to exhibit, when to show component elements of an index, and so on. Moreover, inferences are directed, at least implicitly, by the material selected and the manner of presentation. The reader is left to determine, without explicit guidance, what patterns are present in the data, what they mean, and what importance to attach to this meaning.

Information about the accuracy of the data is essential to proper statistical interpretation, and, as we noted earlier, the reader is not given this information in most instances. Although many of the indicators included are presented in the form of general purpose statistics, it is all the more difficult to discuss accuracy when we don't know what questions we wish our data to help us answer. It would have been helpful if Social Indicators, 1973 had indicated in more detail the purpose of including various charts and data-sets.

Most of the measures reported in Social Indicators, 1973 have been drawn from existing statistical series produced by federal government agencies, and the bulk of these measures are reported in the same manner as in the official reports of the various agencies, or as in the Statistical Abstract. In addition to the measures reported in Social Indicators, 1973, the volume also includes impressive graphical displays. In some cases, such as in the Population chapter, the graphical displays include some statistical projections, which are based upon various assumptions regarding fertility, mortality, and migration. The graphical presentation of this material is highly informative.

We feel that more projections and more detailed statistical analyses are desirable in a social report. In addition, we believe that statistical analyses should be coupled with at least some statistical interpretation and comment.

An examination of the Public Safety chapter will allow us to focus on the type of analysis, interpretation, and comment that is both feasible and desirable. If you glance quickly at Charts 2/1 through 2/3, [; 44-45], and 2/15 through 2/17 [; 53-55], you cannot escape the (naive and perhaps erroneous) conclusion that the rate of criminal offenses has been increasing over time, although there is considerable variability in the rate of increase for different categories of crime. You may even notice the downturn in property crime for 1972, indicated in Charts 2/15 through 2/17. Moreover, Charts 2/13 and 2/14 seem to indicate that the rate of commission of violent crimes (in urban areas) is highest for 15-24 year olds and for Negroes; and Charts 2/8 and 2/9 seem to indicate that the offenders (in a sample from 17 major cities) are mostly male and that they are to a large extent Negro rather than White.

Having drawn these apparent conclusions, we now ask whether they are warranted. If so, then we might ask what are the causes of the increases in crime and the means by which crime can best be prevented or controlled. To answer the first of these questions, we must know something about the error structure of the indices being used. Some of the limitations of the Uniform Crime Reports data are discussed [; 60-61], but little attention is given to accuracy in the statistical sense. There are many who argue that the statistical limitations of the data make it

difficult to determine not only the magnitude but also the direction of changes in rates. Albert Biderman, for example, has stated:

"I contend that most of the sources of error operate to inflate the newer figures relative to the older ones, resulting in a false picture of rapidly increasing lawlessness among the population. With respect to most of these sources of error, it is extremely difficult and sometimes impossible to give quantitative expression to the factor.

Nevertheless, in examining several published criticisms of the index, and in subjecting it to my own critical examination, I believe that the following three conclusions emerge:

1. The errors and biasing factors affecting the Crime Index largely operate to show spurious increases, rather than decreases, in the rate.

2. The Crime Index does not provide a sound basis for determining whether criminal behavior is increasing, or decreasing, in the United States.

3. The Crime Index is highly sensitive to social developments that are almost universally regarded as improvements in the society. Thus, it is altogether possible that year-to-year increases in crime rates may be more indicative of social progress than of social decay."¹¹

Are there other simple explanations for aspects of the apparent increases?

Some explanation is suggested by a reworking of existing data relating to the size and age composition of the population. Since Chart 2/13 reveals that young people commit a disproportionate share of crime, even if the propensity to commit crimes remained constant over time for all age groups, an increasing proportion of individuals in the 15 to 24 year-old bracket could lead to an increasing crime rate.¹² Other

¹¹ Biderman, A. [1966]. "Social Indicators and Goals," pp. 68-153 in Social Indicators, edited by R. A. Bauer, MIT Press, Cambridge, Mass.

The Crime Index referred to in the quotation is a composite index based on the Uniform Crime Reports series for criminal homicide, forcible rape, robbery, aggravated assault, burglary, larceny, and automobile theft.

¹² Another factor having substantial impact on crimes committed by males in the 15-24 year-old age cohort is military service. This point deserves considerable attention, but we do not pursue it here.

demographic characteristics that obviously should be considered in a similar manner are race, sex, and geographical location. There is a simple and straightforward statistical technique called "standardization," well known to demographers and epidemiologists, which adjusts rates or proportions for such factors. While the figures that we need in order to standardize the Uniform Crime Reports rates are not given in Social Indicators, 1973, some relevant direct standardization calculations have been carried out by the Commission on Population Growth and the American Future. They report that

"About 28 percent of the reported increase between 1960 and 1970 in the number of arrests for serious crimes can be attributed to an increase in the percentage of the population under 25. Another 22 percent of the of the increase can be explained by the growing size of the population and other demographic factors. Thus, population change alone accounted for about half of the reported increase in the number of arrests for serious crime over the past decade."¹³

The inclusion of such standardized rates in future editions of Social Indicators would be informative. If such analyses were applied to rates for various types of crimes, we might be able to decide if, for example, the 1972 dip in property crimes is spurious. More important for these future reports are new forms of data such as those now being produced via the National Crime Survey. These new data make possible more thorough statistical analyses. Hopefully, the Law Enforcement Assistance Administration will carry out the appropriate statistical analyses, and will provide OMB with appropriate summaries.

¹³ Population and the American Future, (1972), U. S. Government Printing Office, Washington, D. C., p. 22.

In Social Indicators, 1973 we are given rates per 100,000 population rather than the actual numbers of arrests referred to in this quotation. These reported rates already adjust for the growing size of the population and so standardization could not be expected to account for half of the increase as given in Social Indicators, 1973.

It has been suggested that other factors that may possibly account for the increases in the crime rate over time are:

"... (1) more widespread and intense identification with the norms of the national society, (2) greater integration and effectiveness of the economic and social systems, and (3) more effective operation of the formal agencies of control, such as police and courts."¹⁴

We have yet to come to grips with the possible causes of crime, and the policy implications for its control. What we need, besides more accurate reports on the incidence of crime, are real experiments. We also need measures of potentially related conditions such as the extent and dimensions of narcotics addiction, and also a variety of possible "leading" indicators (as the economists would say) for criminal activity such as, e.g., school truancy rates. Exploring the interrelationships among such variables and their relationship with crime indicators would involve careful statistical analysis.

Other problems in the interpretation of crime statistics revolve around the issue of the incidence of offenses versus the prevalence of offenders, and the use of longitudinal versus cross-sectional data. The Uniform Crime Reports and the victimization data report cross-sectional incidence of crime, but they shed no light on the question of whether certain segments of society, after adjustment for relative size, are committing an increasing or a decreasing amount of crime. The development of an adequate longitudinal

¹⁴Biderman, A. [1966], "Social Indicators and Goals," p. 115-116.

system of criminal statistics, as proposed by the President's Commission on Federal Statistics might shed light on this problem. Again, these are statistical as well as substantive matters.

For a final and somewhat more detailed example of statistical concern, we turn to the Technical Notes for the Public Safety chapter [; 61], where we are told of the concern of crime analysts but not of statistical analysts:

"... a serious problem with the NORC survey was the small sample. Of approximately 2,100 crime incidents identified from interviews carried out in 10,000 households, only 18 were forcible rapes. Crime analysts have questioned the validity of a national rape victimization rate on such a small number of incidents, particularly rates by race and age."

Firstly, we feel that this statement conveys a misconception regarding the accuracy of estimated rates in sample surveys. The standard deviation of an estimated rate decreases as (a) the size of the sample increases, and (b) the size of the true rate decreases (for rates less than 0.5).¹⁵ With respect to the NORC survey, the sample size (10,000) households is relatively large in statistical terms, and the rape rate (estimated at approximately 0.002) is relatively small in statistical terms, so the standard deviation of the estimated rate will be relatively small. Although the 0.002 rate is relatively small in statistical terms, it is not at all small from some other points of view. Accuracy is a relative matter, and if we wished to compare the victimization rate from this survey with a rate produced from other

¹⁵ If p is the true rape rate and n is the sample size, then the standard deviation of the observed rate or proportion is $\sqrt{p(1-p)/n}$. On the other hand, it is of course true that the coefficient of variation (standard deviation divided by the mean which is equal to $\sqrt{p(1-p)/n}/p = \sqrt{(p-1)/n}$, will increase as p decreases toward zero.

sources, the accuracy of the estimated rate may not be high enough to detect small relative differences.

Secondly, the part of the statement dealing with the more detailed breakdowns of crime incidents by race and age seems to ignore the statistical activity of the past decade on the analysis of multiple cross-classifications. By using various statistical techniques, such as unsaturated loglinear models, we can smooth cross-classified data where many of the original cell counts are zero. The resulting smoothed data can then be used to get useful estimated rates.¹⁶

Such statistical methods could be applied to analyze more completely data reported in various chapters of Social Indicators, 1973, e.g., in the Public Safety and Education chapters, but the data in Social Indicators, 1973 itself were not sufficiently detailed for these analyses. In some attempts to go to the source, we found either that the data were from unpublished reports (unavailable to us), or that the cited sources still did not contain the relevant detailed information required. The problem here resides in the limitations of the statistical reports currently produced by some federal agencies and private research organizations which form the basis of Social Indicators, 1973.

¹⁶ For detailed discussion of these techniques see, for example:
 (a) Bishop, Y. M., Fienberg, S. E., and Holland, P. W. [1974]. Discrete Multivariate Analysis: Theory and Practice, MIT Press, Cambridge, Mass.
 (b) Goodman, L. A. [1972]. "A general model for the analysis of surveys, Amer. J. Sociol. 77, 1035-1086.

Summary

In this review of Social Indicators, 1973 we have focused on four major statistical considerations:

- (a) the need for care in reporting data in graphical form,
- (b) the need for detailed discussion of the error structure associated with each of the indicators reported,
- (c) the need for careful statistical analyses of the data presented in the volume and of related data obtained from other sources,
- (d) the need for statistical interpretation of the information reported.

We envision an increasingly important role for statistical analysis and related considerations in future editions of Social Indicators.

Legends for Figures

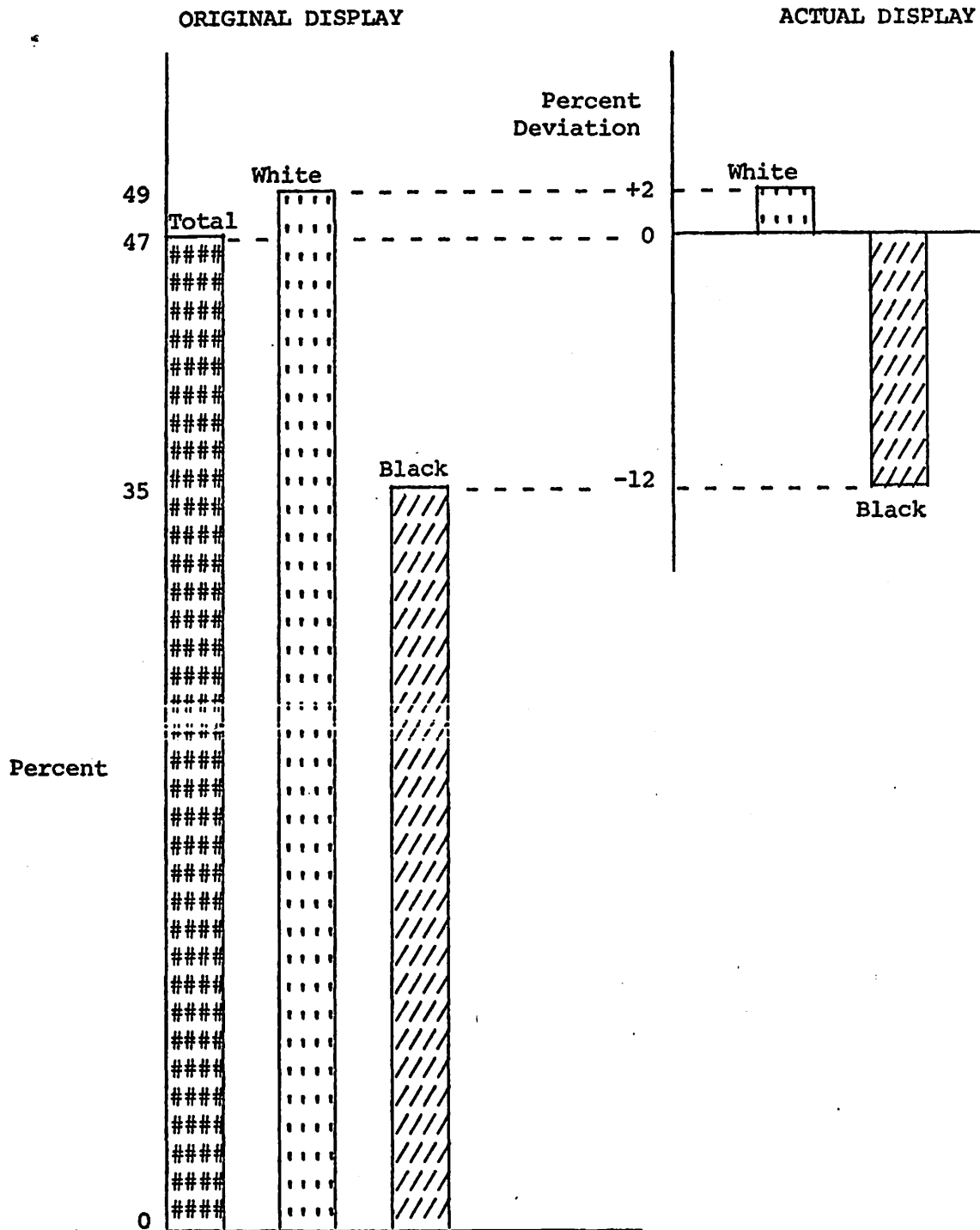
Figure 1. Science Achievement by Color (1970). Actual Display is as in Chart 3/9 of Social Indicators: 1973. Original Display is similar to that used in unpublished report of National Assessment of Educational Progress.

Figure 2. Science Achievement by Color. Similar to Actual Display of Figure 1, but with population size of Black and White groups interchanged.

Figure 3. Science Achievement by Color. The Display here is similar to Original Display of Figure 1, but with bars representing hypothetical interquartile ranges.

Figure 1*

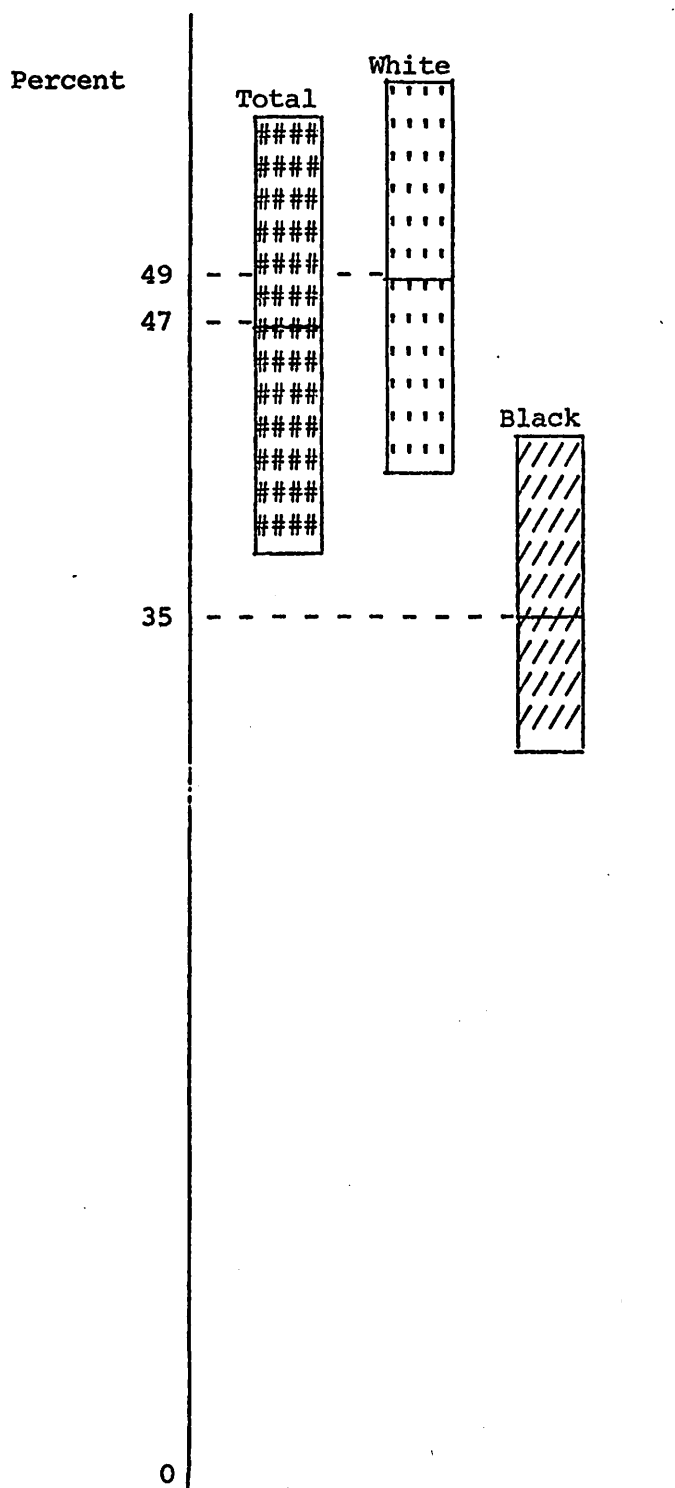
SCIENCE ACHIEVEMENT BY COLOR (1970)



*Actual Display is as in Chart 3/9 of Social Indicators, 1973 [45;83]. Original Display is similar to that used in the unpublished report of the National Assessment of Educational Progress.

Figure 3*

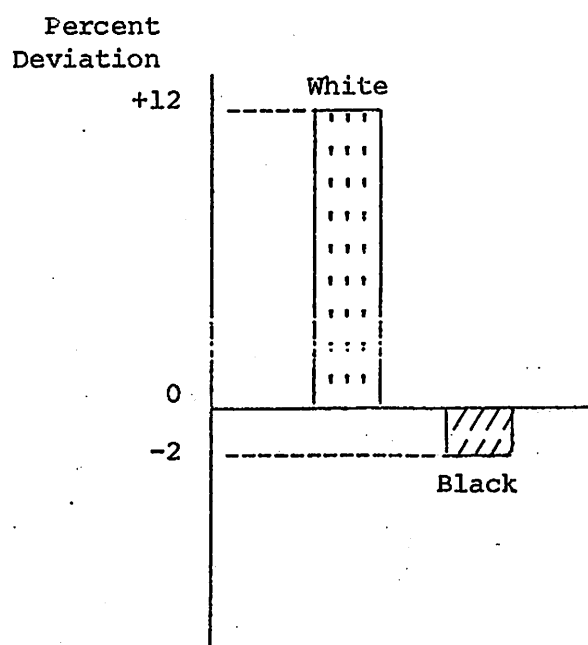
SCIENCE ACHIEVEMENT BY COLOR



*The Display here is similar to Original Display of Figure 1, but with bars representing hypothetical interquartile ranges.

Figure 2*

SCIENCE ACHIEVEMENT BY COLOR



* Similar to Actual Display of Figure 1, but with population size of Black and White groups interchanged.